

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 17:47:40 ; Search time 81 seconds
(without alignments)
1032.779 Million cell updates/sec

Title: US-09-782-587B-1

Sequence: 1 ANAFLLXLRPGSLRXCKXX.....LQKMRSPRPVLLRAPP 406

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues.

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_21:***
2: sp_archaea:***
3: sp_bacteria:***
4: sp_fungi:***
5: sp_human:***
6: sp_invertebrate:***
7: sp_mammal:***
8: sp_mmc:***
9: sp_organelle:***
10: sp_phase:***
11: sp_plant:***
12: sp_rodent:***
13: sp_virus:***
14: sp_vertebrate:***
15: sp_unclassified:***
16: sp_rv1rus:***
17: sp_bacteriophage:***
18: sp_archaeophage:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2161	98.8	701	4	Q96PQ8
2	1558.5	71.3	446	11	Q61109
3	913.5	41.8	433	13	Q90YK1
4	863	39.5	469	6	Q9GMD9
5	849	38.8	461	6	Q9GMD9
6	847	38.7	456	4	Q14316
7	832	38.0	481	11	Q54740
8	830	38.0	482	11	Q63207
9	829	38.0	481	11	Q63207
10	825	37.9	481	11	Q63207
11	825	37.7	461	6	Q95ND6
12	796.5	36.4	456	6	Q9TRD0
13	781	35.7	159	6	Q28510
14	769.5	35.2	460	11	Q91VH8
15	763.5	34.9	460	11	Q91VH8
16	660	30.2	159	6	Q28286

17	477	21.8	638	11	Q8R0P5	Q8R0P5 mus musculus
18	470.5	21.5	624	11	Q9DAT3	Q9DAT3 mus musculus
19	464.5	21.2	624	11	Q91Y47	Q91Y47 mus musculus
20	463.5	21.2	653	11	Q8VCS4	Q8VCS4 mus musculus
21	458	20.9	643	6	Q97506	Q97506 sus scrofa
22	456.5	20.9	799	11	Q9DB10	Q9DB10 mus musculus
23	453	20.7	277	5	Q96899	Q96899 scolopendra
24	446	20.4	339	11	Q99L44	Q99L44 mus musculus
25	441.5	20.2	1089	5	Q873A0	Q873A0 ciona intes
26	431	19.7	399	11	Q9COW3	Q9COW3 mus musculus
27	427	19.5	607	13	Q91001	Q91001 gallus galli
28	425	19.4	237	13	Q91515	Q91515 fugu rubrip
29	424	19.4	492	13	Q96T73	Q96T73 homo sapien
30	422	19.3	608	13	Q9PTW7	Q9PTW7 structio ca.
31	420	19.2	701	11	Q9J3S9	Q9J3S9 ratius norv
32	419.5	19.2	868	5	Q91IV3	Q91IV3 polyanthra
33	419	19.2	467	5	Q967X8	Q967X8 panulirus a
34	418.5	19.1	331	11	Q8R1A6	Q8R1A6 mus musculus
35	418	19.1	624	6	Q95ME7	Q95ME7 oryctolagus
36	417	19.1	787	5	Q9VEY6	Q9VEY6 drosophila
37	416	19.0	812	11	Q9R0W3	Q9R0W3 ratius norv
38	415.5	19.0	855	11	Q9J3J7	Q9J3J7 ratius norv
39	415	19.0	244	13	Q8QGW3	Q8QGW3 anguilla ja
40	415	19.0	812	11	Q91WJ5	Q91WJ5 mus musculus
41	414.5	19.0	247	13	Q9W7O5	Q9W7O5 paralyticth
42	414	18.9	505	5	Q965V4	Q965V4 halocynthia
43	411.5	18.8	250	5	Q9V514	Q9V514 drosophila
44	411	18.8	490	11	Q920K3	Q920K3 ratius norv
45	410.5	18.8	267	5	Q9BK47	Q9BK47 ludia foli

ALIGNMENTS

RESULT 1	ID	SEQUENCE FROM N.A.	PRELIMINARY:	PRT:	701 AA.
Q96PQ8	Q96PQ8	01-DEC-2001 (TREMBLrel. 19, Created)			
Q96PQ8	Q96PQ8	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
Q96PQ8	Q96PQ8	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
Q96PQ8	Q96PQ8	Factor VII active site mutant Immunocongulate.			
Q96PQ8	Q96PQ8	Homo sapiens (Human).			
Q96PQ8	Q96PQ8	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Q96PQ8	Q96PQ8	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
Q96PQ8	Q96PQ8	NCBI_TaxID=9606;			
Q96PQ8	Q96PQ8	[1]			
Q96PQ8	Q96PQ8	SEQUENCE FROM N.A.			
Q96PQ8	Q96PQ8	MEDLINE=21477448; PubMed=11593034;			
Q96PQ8	Q96PQ8	Hu Z., Garen A.;			
Q96PQ8	Q96PQ8	"Targeting tissue factor on tumor vascular endothelial cells and tumor			
Q96PQ8	Q96PQ8	cells for immunotherapy in mouse models of prostatic cancer.";			
Q96PQ8	Q96PQ8	Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).			
Q96PQ8	Q96PQ8	EMBL: AF272774; AAK58686.1; -			
Q96PQ8	Q96PQ8	InterPro: IPR000152; Asx_hydroxyl.			
Q96PQ8	Q96PQ8	InterPro: IPR000561; EGF-like.			
Q96PQ8	Q96PQ8	InterPro: IPR000742; EGF_2.			
Q96PQ8	Q96PQ8	InterPro: IPR001881; EGF-Ca.			
Q96PQ8	Q96PQ8	InterPro: IPR003006; Ig_MHC.			
Q96PQ8	Q96PQ8	InterPro: IPR001254; Ser_protease_Try.			
Q96PQ8	Q96PQ8	InterPro: IPR000294; Vitk_dep-Gla.			
Q96PQ8	Q96PQ8	Pfam: PF00008; EGF_2.			
Q96PQ8	Q96PQ8	Pfam: PF00594; gla; 1.			
Q96PQ8	Q96PQ8	Pfam: PF00047; Ig_2.			
Q96PQ8	Q96PQ8	Pfam: PF00089; trypsin; 1.			
Q96PQ8	Q96PQ8	SMART: SM00181; EGF_2.			
Q96PQ8	Q96PQ8	PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.			
Q96PQ8	Q96PQ8	PROSITE: PS00022; EGF_1; UNKNOWN_1.			
Q96PQ8	Q96PQ8	PROSITE: PS01186; EGF_2; UNKNOWN_1.			
Q96PQ8	Q96PQ8	PROSITE: PS01187; EGF_CA; UNKNOWN_1.			
Q96PQ8	Q96PQ8	PROSITE: PS00011; GLU_CARBOXYLATION; UNKNOWN_1.			
Q96PQ8	Q96PQ8	PROSITE: PS00290; IG_MHC; UNKNOWN_1.			
Q96PQ8	Q96PQ8	PROSITE: PS50240; TRYPSIN_DOM; 1.			

Phan T., Thankavel B., Jagadeeswaran P.;
 "Demonstration of the extrinsic coagulation pathway in teleostei:
 RT Identification of zebrafish coagulation factor VII."
 RT Proc. Natl. Acad. Sci. U.S.A. 98:8768-8773(2001).
 DR EMBL: AY040345; AAK74192.1;
 DR MEROPS: S01.215;
 DR InterPro: IPR000561; EGF_1like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR InterPro: IPR000294; VitK_dep_GLA.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00594; gla_1.
 DR Pfam: PF00089; trypsin_1.
 DR SMART: SM00179; EGF_CA_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
 DR PROSITE: PS00011; GLU CARBOXYLATION; UNKNOWN_1.
 DR PROSITE: PS00240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER; UNKNOWN_1.
 DR HydroLase; Serine protease.
 KW SEQUENCE 433 AA; 48680 MW; CD9D1B179601BA4C CRC64;

Query Match 41.8%; Score 913.5; DB 13; Length 433;
 Best Local Similarity 45.9%; Pred. No. 8.6e-81;
 Matches 178; Conservative 59; Mismatches 138; Indels 13; Gaps 6;

8 LKPGSLKRXKXKXQCFXXARXIFKDAAXRTKLFMISDGDCCASSPCONGSGCKXQ-LQ 66
 46 LKGNREKCLEKCEYEAREVEFEHTNEATNEFWKLYDVADHCASSPCHEHGLCTQNMAD 105
 67 STICFLPAFEGNCTHKKDQ-LICVNGGCEYOYCSDHGTGRSCRCHEGYSLLADGV 125
 106 SYMCLAPGFSFGHCEQSIGDVLDSCLHNGCEHCTEGDG-RNMSCADGYLDSMGQ 164
 126 SCTPTVEYPCGKIPLEKRNASKPQ-----GRIVGKVCPRGEEPPWVLLVNGAQLCGGT 181
 165 KRSHEVEPCGKVPRLQAKADHVDLSRIYSGSECPGKHPWVLLLYGEGKCGGV 224
 182 LINTIVVSAACFDKIKMMRNLIIVLGEHDLSEHGDESRVAOVIIIPSTVPGTNR 241
 225 IKPFWILTAHCELEK-KVFLRIYAGEHDLVDEGTEDLIQVDQFTHPAVSEIADS 283
 242 DIALLRLHQPVLVDHVPDLCPERTFSERTLAFAVFSLVSGWGLD3GATALELVLN 301
 284 DIALLRLRPIYVAVPVCLPLREMARELAWSKHYSGMKRSEDEGPTSRLLRLL 343
 302 VRLMTPDCLQOSRKVGDSPNITEYFPCAGYSKDSKDSKSGSGGPHAHYRGTYLTGI 361
 344 VRIITQECVQVS-----NLTLTSMFCAGYIEGRDSCKDSGSGPLVRYRDTAFLLGI 398
 362 VSMGOCATVGHFGVYTRVSOYIEMLOK 389
 399 VSMGKCARPSGSGITRYNTLQWIRQ 426

RESULT 4
 09GMD9 PRELIMINARY: PRT: 469 AA.

01-MAR-2001 (TREMBLrel. 16, Created)
 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Coagulation factor X.
 OS Ornithorhynchus anatinus (Duckbill platypus).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
 NCBI_TaxID=9258;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21015017; PubMed=11132153;
 Pootafshar M., Aveskog M., Munday B., Hellman L.;

"Identification and structural analysis of four serine proteases in a
 RT monotreme, the platypus, Ornithorhynchus anatinus."
 RT Immunogenetics 52:19-28(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL: AF275654; AAC00453.1;
 DR HSSP: P00742; 1XB.
 DR MEROPS: S01.216;
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF_1like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR002383; GLA blood.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR InterPro: IPR000294; VitK_dep_GLA.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00594; gla_1.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00001; GLABLOOD.
 DR SMART: SM00181; EGF_2.
 DR SMART: SM00179; EGF_CA_1.
 DR SMART: SM00001; EGF_1like; 2.
 DR SMART: SM00069; GLA_1.
 DR PROSITE: PS00020; TRYP_SPC; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA_1.
 DR PROSITE: PS00011; GLU CARBOXYLATION; 1.
 DR PROSITE: PS00011; GLU CARBOXYLATION; 1.
 DR PROSITE: PS00240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 KW HydroLase; Serine protease.
 KW SEQUENCE 469 AA; 52196 MW; 4C66C230D0758F6A CRC64;

Query Match 39.5%; Score 863; DB 6; Length 469;
 Best Local Similarity 38.4%; Pred. No. 8.3e-76;
 Matches 168; Conservative 72; Mismatches 143; Indels 54; Gaps 7;

1 ANAFLLXLRPSKXKXKXQCFXXARXIFKDAAXRTKLFMISDGDCCASSPCONGSG 60
 41 ANSLFEBELKGNLEBCENETGSEARREVEFDTDNEMWNLYKNGCCEIOPCNOGY 100
 61 CKQLOSLYICFLPAEGRNCTHKKDQ-LICVNGGCEYOYCSDHGTGRSCRCHEGYSL 120
 101 CKDGLAEYTCCLASAGEYEGKNCIDT--TTVKISLNNGDCEOFCKSVNTV-VCSAOGYIL 157
 121 LADVGCTPTVEYPCGKIP-LERNAKSPQ----- 151
 158 GDDQKSCIPVPPPCGKLVGRKRSBELPEODGDNAAHVAEDYLEATENPEAEPPDN 217
 152 -----RIVGKVCPRGEEPPWVLL--VNGAQLCGTLLINTIVVSAACFDK 197
 218 TLAEPGENALVIVGREGHOECPQALLVNDENGOGFCGTTINLEYIILSAACHMQ 277
 198 IKMNRLLIIVLGEHDLSEHGDESRVAOVIIIPSTVPGTNRHDIALLRLHQPVLVDH 257
 278 AKRFX--VAVGRDDEKKDSSMAHEVEKVIHYSKFKVYDQFDIAVILKTPITFRMN 334
 258 VVPLCLPERTFSERTLAFAVFSLVSGWGLD3GATALELVLNVRMTPOCLQOSRV 317
 335 VSPACLPEDKADILMKNKAGVSGFGRVHEGRSTVLKMLEVYVERTCKOSS-- 391
 318 GDSNITEYFPCAGYSDGSKDSDGSGGPHAHYRGTYLTGIIVSMGOCATVGHFGVY 377
 392 --SFDITPMFACAGYSRPEDAQGDSDGSGPHTKYNDYTFYVGIYVSGGCAQNGKFGY 449
 378 TRVSQYIEMLOKLMSE 394
 450 TKAATFLSWIKRMROK 466

ID	Q95ND7	PRELIMINARY	PRT	461 AA.
AC	Q95ND7			
DT	01-DEC-2001	(TREMBLrel. 19, Created)		
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)		
DE	Coagulation factor XI.			
GN	P9.			
OS	Pan troglodytes (Chimpanzee).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.			
OX	NCBI_TaxID=9598;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=504;			
SA	Saltu Y.;			
RT	"Comparison of DNA and protein polymorphisms between humans and			
RT	chimpanzees.";			
RL	Genes Genet. Syst. 0.0-0.0(2001).			
DR	EMBL: AB062470; BAB58885.1; JOINED.			
DR	EMBL: AB062458; BAB58885.1; JOINED.			
DR	EMBL: AB062460; BAB58885.1; JOINED.			
DR	EMBL: AB062462; BAB58885.1; JOINED.			
DR	EMBL: AB062464; BAB58885.1; JOINED.			
DR	EMBL: AB062466; BAB58885.1; JOINED.			
DR	EMBL: AB062468; BAB58885.1; JOINED.			
DR	InterPro: IPR000152; Asx_hydroxyl.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR000742; EGF-2.			
DR	InterPro: IPR001881; EGF_Ca.			
DR	InterPro: IPR001234; Ser. protease_Try.			
DR	InterPro: IPR000294; VitK_dep_GLA.			
DR	pfam: PF00008; EGF_2.			
DR	pfam: PF00594; gla; 1.			
DR	pfam: PF00089; trypsin; 1.			
DR	PROSITE: PS00010; Asx_HYDROXYL; UNKNOWN_1.			
DR	PROSITE: PS00022; EGF_1; UNKNOWN_1.			
DR	PROSITE: PS01186; EGF_2; UNKNOWN_2.			
DR	PROSITE: PS01187; EGF_Ca; UNKNOWN_1.			
DR	PROSITE: PS00011; GLU_CARBOXYLATION; UNKNOWN_1.			
DR	PROSITE: PS00240; TRYPsin_DOM; 1.			
DR	PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.			
DR	PROSITE: PS00135; TRYPsin_SER; UNKNOWN_1.			
KW	Hydrolase; Serine protease.			
SO	SEQUENCE 461 AA; 51764 MW; 30628575COF77F45 CRC64;			
Query Match	38.8%; Score 849; DB 6; Length 461;			
Best Local Similarity	39.3%; Pred. No. 1.9e-74;			
Matches 163; Conservative 70; Mismatches 130; Indels 52; Gaps 8				
QY	11 GSLRXRCXXXCSPFXARXIFKDXARTLFLNISYSDGQCASSPQNGSGCKDQLOSTIC 70			
Db	58 GNLRRECHKEECSBEAREVENTERTTEFKQYVDGDCSBNPOLNGSGCKDDINSYEC 117			
QY	71 FCLPAFGRNCETHKDDQLICVNGGCEQYCSDHGTGRKSCRCHEGYSILADVSCTPT 130			
Db	118 WCPGFGEFGKNCEL----DYTCNINNGRCGECKNSADKKVYCSCTEGIRLLENQSCPEA 173			
QY	131 VEYCCGILPILEKRNASKPQ-----GRIVGKVC 159			
Db	174 VPPFCGRVSVQTSKLRRAETVPPDVYVNSTEATITLDTIGSTQSFNDPTRYVGCEDA 233			
QY	160 PKGCEPQOVLLVNGAOLCGTTLNTIMVVSAAAGCFK-IKNMNLIVLGEHDSLEND 218			
Db	234 KPGGFPQOVVLLNGKVDVAFCGGSIVNEKRIYTAACVDIVGK----ITVAAEHNIEFEH 289			
QY	219 DEQSRARVQYIISTYVPGTT--NHDIALLRQPVVLTIDHVVLELCPERTFSRTIAFV 276			
Db	290 TEQRRNVIRIIRPHNYNAINKYNDHIALLELDEPVLNYSVPTCIADK---EYTNIFL 346			
QY	277 RF--SLVSGMGQLDRGATLLELVAVLWVPLMTDCCLOQSKKVDSPNITEYMFCAGYSD 334			

Qy	335	GSKSCGSDGGGPHATYRGCTWLTGTVSGCCCATYGHGVYTRVSYQIEWLQK	389
Db	347	KFGSGYGVSGMGFRFHKGRSLVLYQLYLRVPLVDRACTCRKFK-----TIVNNFCAGFHE	401
Db	402	GGRRSCGGSDGGGPHATVEGTSFTLTGIIWGCACMKGYITKVSRYVNWIK	456

```

OY 11 GSLAKRXXKXCCSFYXARXIFDAXRTKLFWISYSDGDCASSPCONGSCKRQLOSLYIC 70
DB 53 GULEHECKEKSFEFARAEVFENTERTEFMKQYVDDCESNPLCINGSCSKDINSYEC 112
OY 71 FCLPAFEGNCEETHKDDOLICVNEGCEOYCSDHGTGRSCRCHEGYSILADGVCSTPL 130
DB 113 MCPFEFGNCEL----DYTCNIKGRRCFOFCKNSADNKVCSCTGTYRLAEOKCEPA 168
OY 131 VEYPCGKPILEKRNASKPO-----GRIVGKYC 159
DB 169 VPFPGKRVSVQTSKLTAEVFPDVTNSTEATILDNITQSTQSFNDFETRVGCEDA 228
OY 160 PKGECFMOVLVLVNGAGGLINTIWMVSAHCDKIKNNRNLIAYLGEHDSHDD 219
DB 229 KPGCFPMOVLVNGKVDAPFCGGSIVNEKMTVTAHC---VETGVKTIIVAGENHIEETHT 285
OY 220 EQRRAVAVIIPSTVPGTT--NHDIALLRHQPVLTDPVYVPLCLPEFTSERTLAFFR 277
DB 286 EOKRVAVIILPHNNVNAIINKNHDIALLEDEPLVNSYVPICIADR---EYVITFLK 342
OY 278 F--SLVSGMGLDLBGATALELMLVNPRLMTQDCLQOSRKVGDSPNITEYMFCAYSOG 335
DB 343 FGSGVSGMGRVFNHGRSALVLYQYLRVPLVDRATCLRSTKF-----TIVNMECAGFHG 397
OY 336 SKDSCKDGGPHTATYRGTYLTGIVSGGCCATVGHGVYTRVSYITELQK 389
DB 398 GRDSCGSGGPHVTEVESTFLTGISMGECAMKKGKGYTKVSRYVNMKE 451

RESULT 7
OY 054740 PRELIMINARY: PRT: 481 AA.
AC 054740.
DB 01-JUN-1998 (TREMBLrel. 06, Created)
DB 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DB 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DB Coagulation factor X precursor (EC 3.4.21.6).
DB F10 OR FA10.
DB Mus musculus (Mouse).
DB Plasmid paluascript.
DB Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
DB Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
DB NCBI_Taxid=10090;
DB [1]
DB SEQUENCE FROM N.A.
DB TISSUE=LIVER.
DB MEDLINE=9845493; PubMed=9783672;
DB Heldmann H.H., Kontermann R.E.;
DB "Cloning and recombinant expression of mouse coagulation factor X.",
DB Thromb. Res. 92:33-41(1998).
DB -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
DB TRYPSIN FAMILY.
DB EMBL: AJ222677; CAAL0933.1;
DB HSSP: P00742; IYKA.
DB MEROPS: S01.216;
DB MGD: MGI:103107; F10.
DB InterPro: IPR000152; Asx_hydroxyl.
DB InterPro: IPR001314; Chymotrypsin.
DB InterPro: IPR000561; EGF-like.
DB InterPro: IPR000742; EGF 2.
DB InterPro: IPR001881; EGF_Ca.
DB InterPro: IPR002383; GLA_Dlood.
DB InterPro: IPR001254; Ser-protease_Try.
DB InterPro: IPR000294; VitK_dep_GLA.
DB Pfam: PF00098; EGF_2.
DB Pfam: PF00594; gla; 1.
DB Pfam: PF00089; trypsin; 1.
DB PRINTS: PR00072; CHYMOTRYPSIN.
DB PRINTS: PR00001; GLABLOOD.
DB SMART: SM00179; EGF_CA; 1.
DB SMART: SM00001; EGF_like; 1.
DB SMART: SM00069; GLA; 1.
DB SMART: SM00020; Tryp_Spc; 1.

```

```

DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE: PS00240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR Calcium-binding; EGF-like domain; glycoprotein; Hydrolase; plasmid;
DR Repeat; Serine protease; signal.
DR SIGNL 1 40 POTENTIAL.
FT CHAIN 41 481 COAGULATION FACTOR X.
SQ SEQUENCE 481 AA; 53986 MW; CF702DE5ER9D97AE CRC64;

Query Match 38.0%; Score 832; DB 11; Length 481;
Best local Similarity 36.8%; Pred. No. 9.3e-73;
Matches 162; Conservative 82; Mismatches 142; Indels 54; Gaps 8;

OY 1 ANAFLLXLRPSGLKRXKXCCSFYXARXIFDAXRTKLFWISYSDGDCASSPCONGS 60
DB 41 ANSFEEFKGNLERECMEICSYEVRREIFEDDEKTRVTKYKDGDCSSPCONGA 100
OY 61 CDDQLOSLYICFCLPAFEGNCEETHKDDOLICVNEGCEOYCSDHGTGRSCRCHEGYSL 120
DB 101 CRDGIQGYCTCSEGFEGNCELFLV--RKLCLRLDNGDCDQFCREONSV--VCSASGYFL 157
OY 121 LADVSCPTVEYPCGKI-----PL-LEKR 144
DB 158 GNDGKSCISTAPFPDPCGKITGTRRRKRSVALINSDSELDLALDDEFLSPENPIELNL 217
OY 145 NASKPOG-----RVGKVCPCGECPMOVLVL--VNGALCGGTLINTIWMVSAHCFDK 197
DB 218 NETOPERSDDLVRIVGECCKDGECPMALLINDNGFCGTLINFEYITLTAHCHQ 277
OY 198 INNNRNLIAYLGEHDSHDDGOSRRVAQVITPSTVPGTTNHDIALLRHQPVLTDR 257
DB 278 ARRFK--VRGDRNTEKEDGNELVHEVDVYIKHKKFORDYDYDIAYLRILKTPITFRMN 334
OY 258 VVPLCLPERTFSERTLAFFRSFLVSGMGLDLRGATALELMLVNPRLMTQDCLQOSKV 317
DB 335 VAPALPDKMAESTLMQKGIYSGFGRHKGQSNILKLEVPYDRNTC-----KL 389
OY 318 GDSNITEYMFCAGYSKDSCKDSCGSPHATYRGTYLTGIVSGGCCATVGHGVCY 377
DB 390 STFSITQNMFCAGYEALLEDACDGSQSPHYTRKNYTYVIGVSGECARAGKGYIT 449
OY 378 TRVSQIYELQKLMSRPP 397
DB 450 TKVTFELKVIDRSMKARVGP 469

RESULT 8
OY 063207 PRELIMINARY: PRT: 482 AA.
AC 063207.
DB 01-NOV-1996 (TREMBLrel. 01, Created)
DB 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DB 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DB Factor X.
DB Rattus norvegicus (Rat).
DB Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
DB Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
DB NCBI_Taxid=10116;
DB [1]
DB SEQUENCE FROM N.A.
DB STRAIN=SPRAGUE-DAWLEY;
DB MEDLINE=96093366; PubMed=8578539;
DB Stanton C., Ross R.P., Hutson S., Wallin R.;
DB "Evidence for competition between vitamin K-dependent clotting factors
DB for intracellular processing by the vitamin K-dependent gamma-
DB carboxylase.",
DB Thromb. Res. 80:63-73(1995).
DB -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE

```

CC TRYPsin FAMILY.
 DR EMBL: X79807; CAA56202.1; -
 DR HSSP: P00742; IYKA.
 DR MEROPS: S01.216; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF-2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR002383; GLA_blood.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR InterPro: IPR000294; Vitk_dep_GLA.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00594; gla; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPsin.
 DR PRINTS: PR00001; GLABLOOD.
 DR SMART: SM00179; EGF_CA; 1.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00069; GLA; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
 DR PROSITE: PS00240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase.
 DR Hydroxylation; Repeat; Serine protease.
 KW SEQUENCE 482 AA; 54265 MW; 0284678E3954A698 CRC64;

Query Match 38.0%; Score 831.5; DB 11; Length 482;
 Best Local Similarity 36.1%; Pred. No. 1e-72;
 Matches 159; Conservative 83; Mismatches 144; Indels 55; Gaps 7;

QY 1 ANAFLXLRPGSLXKXCKXCCSFXXARIFKDAKRTLFWISYSDGQCCASSPCQNGS 60
 DB 41 ANSFEEIKKGNLEECVCEICSEFEAREVEDNEKTEFEWKKYEDGQCCSSPCQNGE 100
 QY 61 CKDLOSTYICFLPAFEGRNCEETHKDDQLICVNEGGEQVCSHDTGKRSRCHEGYSL 120
 DB 101 CRDGLGTYCTCTCTEEFEGKNCLEFY--RKLSLDNGDCDQFCREQNSV-VCSKKGIFL 157
 QY 121 LADGVCSTPVEYPCGKI-----PILEKR 144
 DB 158 GNDGKSCSLTAPEPCGKTKNGRAKRSVALNTNSRPDEDLMPDADILYPTESPSELLNL 217
 QY 145 NASKPQG-----RIVGKVCPCKGCPCMOVLLVNGAO--LCGGLINTIVVSAHCFD 196
 DB 218 NKTPEANSDDVIRIVGQCECKRGCECPWQALLFSDSETDGCGLINEFYILTAHCLH 277
 QY 197 KIKMNRNLIAVLGEHDLSEHGDDEQSRVAQVILIPSTVPPTNHDLILHQQVVLTD 256
 DB 278 QAKRK---VAVGDLNTEQEDGGEVHVEDMILKKNKQORDYDIDLMLRKTPITRE 334
 QY 257 HAVPLCLPEPTFSERTLAFAVPSLSVSGMOLLDRGATLLELVILNVPRLMTQDCIQGSRK 316
 DB 335 NVAPACLPOKDMAEATLMTOKTGIVSGFRTHKROSKVLAEMEVPIVDNNTC-----R 389
 QY 317 VQDSPNTEYMFCAQSYSGSKDSCGSGGPHATYKGTWTLGLIVSGGCAVGHGV 376
 DB 390 LSTFSITQNMFCAGYDAKQBDACGSGGPHVTFKDTYFTVGTGIVSGEGCARKGKGYI 449
 QY 377 YTRVSOYIEMLOKLRSEPR 397
 DB 450 YTKVTAFLKWDIRSKARVGP 470

RESULT 9
 088947 PRELIMINARY; PRT; 481 AA.

AC 088947;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Coagulation factor X precursor.
 GN F10.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL6 X CBA; TISSUE=LIVER;
 RA MEDLINE=98347933; PubMed=9684791;
 RX Liang Z., Cooper A., Deford M.E., Carmeliet P., Collen D.,
 RA Castellino F.J., Rosen E.D.;
 RT "Cloning and characterization of a cDNA encoding murine coagulation
 factor X."
 RL Thromb. Haemost. 80:87-91(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129S1;
 RA Cooper A., Liang Z., Castellino F.J., Rosen E.D.;
 RT "Cloning and Characterization of the Murine Factor X Gene";
 RL Thromb. Haemost. 0:0-0(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPsin FAMILY.
 DR EMBL: AF087644; AAC36345.1; -
 DR EMBL: AF211347; AAF29980.1; -
 DR HSSP: P00742; IYKA.
 DR MEROPS: S01.216; -
 DR MGD: MGI:103107; F10.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF-2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR002383; GLA_blood.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR InterPro: IPR000294; Vitk_dep_GLA.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00594; gla; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPsin.
 DR PRINTS: PR00001; GLABLOOD.
 DR SMART: SM00179; EGF_CA; 1.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00069; GLA; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
 DR PROSITE: PS00240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
 KW Serine protease; Signal.
 FT SIGNAL 1
 FT CHAIN 1
 SO SEQUENCE 481 AA; 54018 MW; 8AC09DE5EF9D271E CRC64;

Query Match 38.0%; Score 830; DB 11; Length 481;
 Best Local Similarity 36.6%; Pred. No. 1.5e-72;
 Matches 161; Conservative 83; Mismatches 142; Indels 54; Gaps 8;

QY 1 ANAFLXLRPGSLXKXCKXCCSFXXARIFKDAKRTLFWISYSDGQCCASSPCQNGS 60
 DB 41 ANSFEEIKKGNLEECVCEICSEFEAREVEDNEKTEFEWKKYEDGQCCSSPCQNGA 100
 QY 61 CKDLOSTYICFLPAFEGRNCEETHKDDQLICVNEGGEQVCSHDTGKRSRCHEGYSL 120

```

Db 101 CRDGGYCTCTGCEGFEKNCLELV--RKLCRLDNGDCDQFCREONS-VCSGASGYFL 157
QY 121 LAGVSGCTPVEVPCGKI-----PI-LEKR 144
Db 158 GNGKSCISITAPPCCGKITTGRRKRSVALNTSDSELDLEDALLDEDFLSPTEPIELNL 217
QY 145 NASKPQG-----RIVGKVCYCPKGECPMVOYLL-VNGAOLCGGTLINTIWMVSAHCFDK 197
Db 218 NETQPERSSDDLVRIYIGRGCKDECPMALLNEDNEGCGGTLIEFILTAAHCLHQ 277
QY 198 IKMNRLLIAYLGEHDLSEHGDQSRRAVAVIIPSTVPQTTHDIALRLHQPVLTDH 257
Db 278 ARRFK---VAVGDRNTEKEEGNEMVHEVDVVIKHNKFORDTYDYDIAVLKTPITFRMN 334
QY 258 VVPLCLPERFESRTLAFAVRSYSGMQLLDGATALLMVLNVPRMLTODCLQOSRKY 317
Db 335 VAPACLPQKDMASTLMTQGTGIVSGFRTHGKROSNIILKLEVPYVRNTC-----KL 389
QY 318 GDSPTITEVFCAGYSDGSKDSCGSGPHATHYRGTWLTGIVSGGCAVGHGIVY 377
Db 390 STFSITQNMFCAGYAKLEDACQDGSQGPHTYFRKNTYVTVGIVSGECARKGKYGII 449
QY 378 TRVSQYIEMLOKIMRSEPR 397
Db 450 TKVTFELKWMIDRSMKARVGP 469

```

RESULT 10

```

099L32 PRELIMINARY; PRT; 481 AA.
AC 099L32:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE Coagulation factor X.
GN F10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1: SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; BC003877; AA03877.1; -.
DR HSSP; P00742; 1XKA.
DR MEROPS; S01_216; -.
DR MGD; MGI:103107; F10.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF-2.
DR InterPro; IPR001881; EGF-2.
DR InterPro; IPR001438; EGF-11.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitL_dep_GLA.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00584; gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00010; EGFBLD.
DR PRINTS; PR00001; GLABLOOD.
DR SMART; SM00181; EGF_2.
DR SMART; SM00001; EGF-like_2.
DR SMART; SM00020; Tryp_Spc; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 1.

```

```

DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
KW Serine protease.
SQ SEQUENCE 481 AA; 54004 MW; BD88E96C8A0B7E7F CRC64;

```

```

Query Match 37.9%; Score 829; DB 11; Length 481;
Best Local Similarity 36.6%; Pred. No. 1,8e-72;
Matches 161; Conservative 83; Mismatches 142; Indels 54; Gaps 8;

```

```

QY 1 ANAFLLXLRPGLSLRCKXKXQCSFYXAXIIFKDAKRTKLEWISDQGCASSPCONGGS 60
Db 41 ANSFEFEFKGNLERECMEICSYEEVRELFEDDEKTKEMWTYKNGDCESSPCNOGA 100
QY 61 CKQLOSLYICFLPAFEGNCEYHKHKKDQILCVNENGCEQYCSDHGTGTRKSCHEGYSI 120
Db 101 CRDGGYCTCTGCEGFEKNCLELV--RKLCRLDNGDCDQFCREONS-VCSGASGYFL 157
QY 121 LAGVSGCTPVEVPCGKI-----PI-LEKR 144
Db 158 GNGKSCISITAPPCCGKITTGRRKRSVALNTSDSELDLEDALLDEDFLSPTEPIELNL 217
QY 145 NASKPQG-----RIVGKVCYCPKGECPMVOYLL-VNGAOLCGGTLINTIWMVSAHCFDK 197
Db 218 NETQPERSSDDLVRIYIGRGCKDECPMALLNEDNEGCGGTLIEFILTAAHCLHQ 277
QY 198 IKMNRLLIAYLGEHDLSEHGDQSRRAVAVIIPSTVPQTTHDIALRLHQPVLTDH 257
Db 278 ARRFK---VAVGDRNTEKEEGNEMVHEVDVVIKHNKFORDTYDYDIAVLKTPITFRMN 334
QY 258 VVPLCLPERFESRTLAFAVRSYSGMQLLDGATALLMVLNVPRMLTODCLQOSRKY 317
Db 335 VAPACLPQKDMASTLMTQGTGIVSGFRTHGKROSNIILKLEVPYVRNTC-----KL 389
QY 318 GDSPTITEVFCAGYSDGSKDSCGSGPHATHYRGTWLTGIVSGGCAVGHGIVY 377
Db 390 STFSITQNMFCAGYAKLEDACQDGSQGPHTYFRKNTYVTVGIVSGECARKGKYGII 449
QY 378 TRVSQYIEMLOKIMRSEPR 397
Db 450 TKVTFELKWMIDRSMKARVGP 469

```

RESULT 11

```

095ND6 PRELIMINARY; PRT; 461 AA.
AC 095ND6:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Coagulation factor XI.
GN F9.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
OX NCBI_TaxID=9596;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-505;
RA Satia Y.;
RT "Comparison of DNA and protein polymorphisms between humans and
RT chimpanzees.";
RL Genes Genet. Syst. 0:0-0(2001).
DR EMBL; AB062471; BAB58886.1; JOINED.
DR EMBL; AB062459; BAB58886.1; JOINED.
DR EMBL; AB062461; BAB58886.1; JOINED.
DR EMBL; AB062463; BAB58886.1; JOINED.
DR EMBL; AB062465; BAB58886.1; JOINED.
DR EMBL; AB062467; BAB58886.1; JOINED.
DR EMBL; AB062469; BAB58886.1; JOINED.
DR InterPro; IPR000152; Asx_hydroxyl.

```

DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF-2.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR001254; Ser.protease_Try.
 DR InterPro: IPR000294; Vltk_dep_GLA.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00594; gla_1.
 DR Pfam: PF00089; trypsin_1.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
 DR PROSITE: PS01187; EGF_CA; UNKNOWN_1.
 DR PROSITE: PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
 DR PROSITE: PS0240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER; UNKNOWN_1.
 DR Hydrolase: Serine protease.
 KW SEQUENCE 461 AA: 51695 MW: 85FA69A525DF65B5 CR664;

Query Match 37.7%; Score 825; DB 6; Length 461;
 Best Local Similarity 39.3%; Pred. No. 4,2e-72;
 Matches 163; Conservative 68; Mismatches 132; Indels 52; Gaps 9;

OY 11 GSLRXKXKXKXCSFXAXXIFKDAKRTKLEWISTSDGQCASSPCONGSGCKDQASYTC 70
 DB 58 GNLERECMEKCSFEAREVEFEMTERTEFEKQYVDGQCESNCLNGSGCKDDINSYEC 117
 OY 71 FCLPAFEGRNCEETHKDDQICVNEGGEYOYCSHGTGRSCRGHEGSLADGVSCTPT 130
 DB 118 WCPFGFEKNCNEL----DYCNINIKNGRECQCKNSADNKVYSCSTEGRIALNENKSCSPA 173
 OY 131 VEYPCGKPILEKRNASK-----PQG-----RIYGVKVC 159
 DB 174 VPFPCGVRVSVQTSKLTFRATVFPDVEDVNSTEATILDNITQPOSKXXXTRAVGGEDEX 233
 OY 160 PKGSCPMOVLVINGAQLCGGTLINTIWNVAACFPK-IKNENLAVAGEHDLSDHG 218
 DB 234 NPGQPRKRVLVNGVADFACGGSIVENKEMIVTAACVDTGVK---ITVAGEHNIEETEH 289
 OY 219 DEQSRVAOVITPSTVYPGTT--NHDIALLRHOPVVLTDVNPVLCIPERFSRTIAFV 276
 DB 290 TEQKRVNIRIIPHNINNAIKYKMHDALEDEPLVLNSTVYPICTADK---EYTNIFL 346
 OY 277 RF--SLVSGMQLDRGATLALVNLVNPRLMTODCLQOSRKVSDSPNITEYMCAGYSYD 334
 DB 347 KFGSGYVSGMGVRFHKKRSALVQLYLRVPLVDRAATCLARSTK-----TIINNFCAGFHE 401
 OY 335 GSKDSCGDSGGGPRATHTRGTWITGIVSGGCAVYGHFVYTRVSOYIEMLOK 389
 DB 402 GGRDSCGDSGGGPRHVEEGTSFTLTGISWGEBCAMKKGIVTKVSRVYVMIKE 456

RESULT 12

O9TRRO PRELIMINARY: PRT: 456 AA.
 AC O9TRRO;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Protein C precursor.
 GN PROC.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_Taxid=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Leeb T., Kopp T., Deppe A., Breen M., Matls U., Brunnberg L.,
 RA Bregig B.;
 RT "Molecular characterization and chromosomal assignment of the canine
 RT protein C gene."
 RL Mamm. Genome 10:135-139(1999).
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE-99371952; PubMed-1043005;
 RA Leeb T., Pfeiffer I., Kopp T., Deppe A., Bregig B.;
 RT "Analysis of canine protein C gene polymorphisms."
 RT Anim. Genet. 30:237-238(1999).
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPsin FAMILY.
 DR EMBL: AJ001979; CA05126.1;
 DR HSSP: P04070; 1PCU.
 DR MEROPS: S01.218;
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR002383; GLA_blood.
 DR InterPro: IPR001254; Ser.protease_Try.
 DR InterPro: IPR000294; Vltk_dep_GLA.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00594; gla_1.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPsin.
 DR PRINTS: PR00001; GLABLOOD.
 DR SMART: SM00181; EGF_2.
 DR SMART: SM00069; GLA_1.
 DR SMART: SM00020; Tryp_Src_1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
 DR PROSITE: PS0240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase;
 KW Hydroxylation; Repeat; Serine protease; Signal.
 FT SIGNAL 1 42
 FT CHAIN 43 192 PROTEIN C LIGHT CHAIN.
 FT CHAIN 193 194 PROTEIN C CONNECTING DIPEPTIDE.
 FT CHAIN 195 456 PROTEIN C HEAVY CHAIN.
 FT SEQUENCE 456 AA: 50813 MW: 7AD3A8C1C34E59FE CR664;

Query Match 36.4%; Score 796.5; DB 6; Length 456;
 Best Local Similarity 39.7%; Pred. No. 2.6e-69;
 Matches 165; Conservative 74; Mismatches 144; Indels 33; Gaps 11;

OY 1 ANAFLLXLRPGSLRXKXKXCSFXAXXIFKDAKRTKLEWISTSDGQCASSPCONGSGCKDQASYTC 52
 DB 43 ANSFLEIRAGSLERECMEKCSFEAREVEFEMTERTEFEKQYVDGQCESNCLNGSGCKDDINSYEC 102
 OY 53 SPONGSGCKDQASYTCFCLPAFEGRNCEETHKDDQICVNEGGEYOYCSHGTGRSCRGHEGSLADGVSCTPT 112
 DB 103 SPCGHSICIDGIGAFHCDGCRGMEGRFCO-HEVSYINCLDNGGSHYCLEEG-GRHC 160
 OY 113 RCHEGYSLADGVSCTPTVYEPGKI-PILEKRN-----NASKPQGITGVKVCYCPKG 162
 DB 161 SCAPGRLGDGDDHLDQCPAVKFPKGRGKQKMKRKLKRDNTDIDIPVLVNGKTVRK 220
 OY 163 ECPQVLLVINGAQL-CGGTLINTIWNVAACFPDKIKMNRNLAVAGEHDLSDHGDDQ 221
 DB 221 ESPMOVVLDSKKKLAGAVLHITSVTLRAHMEDESK---KLIVRLGETDLRRWKEGN 277
 OY 222 SRVAVOVIIPSTVYPGTTNHDIALLRHOPVVLTDVNPVLCIPERFSRTIAFV-RFSL 280
 DB 278 DVDIKEVLLHPNYSKSTNDIALHLADPAIFRSQIVPICLPDSGLARELTQVQGEIV 337
 OY 281 VSGMQLD--RGATLLELVNLVNPRLMTODCLQOSRKVSDSPNITEYMCAGYSYD 338
 DB 338 VTGWRSTKTRNRFV-LNFTINIPAPNNECIOAYNN-----ISENNLCAGIILDSRD 391
 OY 339 SCKDSCGSGPRATHTRGTWITGIVSGGCAVYGHFVYTRVSOYIEMLOKMRSE 394
 DB 392 ACEGDSGGMVTSFNGTWLVLVLSWGEBCGRLHNYGIVTKVSRVLDWIHSIRGE 447

Query	Local	Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
1	ANAFLLXLRPGGLXRXKXKXOCSPFXAXRIFKDXAPRKLMISYSDGOCAS	35.2%	769.5	DB 11	460	158	79	152	41	11
42	ANSFLEEMRGRSLERBCEMEICDPFEAOETQONVEDTLARIKTFDGDSCAPPLDHOD	36.7%	769.5	DB 11	460	158	79	152	41	11
53	SPQONSGCKDLOSLYICFCLEPAEGRNCEHTKDDOLICVNGGCEQYCSHGTGRSC	36.7%	769.5	DB 11	460	158	79	152	41	11
102	SPCCGHGTCIDIGSFSCSCDKGMEGRKFCQOELPFQOCRVN-NGCCLHYCLEESNCR-C	36.7%	769.5	DB 11	460	158	79	152	41	11
113	RCHEGSLIADSVCTPVEYPCGIR-ILEKRNASK-----PQGRVYGKVCIP	36.7%	769.5	DB 11	460	158	79	152	41	11
160	ACAPGTEIADDMRCKSTVFPCCGLGMEIKRRIILKRPDLEDELEPPRIVNTLTK	36.7%	769.5	DB 11	460	158	79	152	41	11
161	KCEBPQVLYLVNLAQOL-CGGTLINTIVWYSAACFPKIKMRNLIAVSEHDSHOD	36.7%	769.5	DB 11	460	158	79	152	41	11
220	OGDSPQOAILDSSKLLAGGVLIHTSVLNAHC---VGTCKLTVRLGEYDLRRRDHW	36.7%	769.5	DB 11	460	158	79	152	41	11
220	EOSRRVQVYITSTVPGTTHDIALRLHDPVYLDLPERTSERTLAFVRS	36.7%	769.5	DB 11	460	158	79	152	41	11
277	EEDLDIKELVHPNTRSSNDIALRLAPATLSKTIYPICLPNNGLAOELTOAGCET	36.7%	769.5	DB 11	460	158	79	152	41	11
280	IYSGMGOLD-----RGATLDELVLVPLMLTODLQGSRVGDSPTTEYFCAGS	36.7%	769.5	DB 11	460	158	79	152	41	11
337	VYTGKQYSDRKDKDRNRRTET-LFTIIPPLVANNEVEVMKVV-----VSEMLDAGIT	36.7%	769.5	DB 11	460	158	79	152	41	11
334	DSKSCGKSGGPHATYRGTYWYLTGIVSGOCATVGFHYTRYVQYIEMLOKMS	36.7%	769.5	DB 11	460	158	79	152	41	11
391	GCTRPACODSDSGPMVFEFGRTWELVGLVSGECGHTNMYGYITKGSYLKMHISYCE	36.7%	769.5	DB 11	460	158	79	152	41	11
394	EPRPGLLRA 403	36.7%	769.5	DB 11	460	158	79	152	41	11
451	K--GVSLKS 457	36.7%	769.5	DB 11	460	158	79	152	41	11

RP SEQUENCE FROM N.A.
RC STRAIN-C57BL;
RA Korf I.;

RT "Complete sequence of UC72A01."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.

DR EMBL: AF318182; AK07918.1; -

DR HSSP: P04070; IPCU.

DR MEROPS: S01.218; -

DR MGD: MG1.97771; PROC.

DR InterPro: IPR000152; Asx_hydroxyl.

DR InterPro: IPR001314; Chymotrypsin.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR001881; EGF Ca.

DR InterPro: IPR002383; GLA_blood.

DR InterPro: IPR001254; Ser_protease_Try.

DR InterPro: IPR000294; VitK_dep_GLA.

DR Pfam: PF00008; EGF; 2.

DR Pfam: PF00594; gla; 1.

DR Pfam: PF00089; trypsin; 1.

DR PRINTS: PR00722; CHYMOTRYPSIN.

DR PRINTS: PR00001; GLABLOOD.

DR SMART: SM00181; EGF; 2.

DR SMART: SM00069; EGF-like; 2.

DR SMART: SM00020; Tryp_Spc; 1.

DR PROSITE: PS00010; ASX_HYDROXYL; 1.

DR PROSITE: PS00022; EGF_1; UNKNOWN_1.

DR PROSITE: PS01186; EGF_2; 2.

DR PROSITE: PS01187; EGF_Ca; 1.

DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.

DR PROSITE: PS50240; TRYPSIN_DOM; 1.

DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE: PS00135; TRYPSIN_SER; 1.

DR Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase;

KW Hydroxylation; Repeat; Serine protease.

SEQUENCE 460 AA: 51784 MW: 02938C25E9D3ED16 CRC54;

Query Match 34.9%; Score 763.5; DB 11; Length 460;

Best Local: Similarity 36.5%; Pred. No. 4.4e-66;

Matches 157; Conservative 79; Mismatches 153; Indels 41; Gaps 11;

QY 1 ANAFLXLRPGSLKXCKXOCSPFXRXRIFPDAXRTKLFMTSYSDGDCAS-----52

DB 42 ANSFLEMRPGSLRECHMEICDLEAEQEIFQNVEDTLAFWIKYFDGDCSAPPLDHOC 101

QY 53 SPCONGSCQDQLOSYICFLPAFEGNCEETHKDDQLICVNGGCEQYCSDHGTGRSC 112

DB 102 SPCGCHGTGIDIGSFCSDKMGWEGKFCQQLRQDCRVN-NGCCLHYCLEESNGRR-C 159

QY 113 RCHEGYSLADGVSTPVEYPCGKIP--ILEKRNASK-----PGGRIVGKVCYCP 160

DB 160 ACAPGYELADDMHMRCKSTVNFPCGKLGMRMKRKRILKRDYDLEDELEPPDPRIYNGTLTK 219

QY 161 KGECPWVLLVNGAQL-CGGTLINTIWNVSAHCFDKIKMRNLINAVLGHDSEHGD 219

DB 220 QGDSFWQAILLDSKKKLACGVLIHTSWLTPAHC---VEGTKLYRLGELYDLRRDHW 276

QY 220 EQSRRAVQIIPSTYVPGTTHNDIALRLHQPVLVDHVPVLCLEPRTSSERTLAFVRS 279

DB 277 ELDDLKIELIHPNVTBSSNDIALRLAPATLSKTIYVICLPNNGLAQELTOAGQET 336

QY 280 LVSGWGLLD-----RGATLELMVNLVPRMLTQDCLQOSRKYGDSPNITEYMFCAYS 333

DB 337 VVTGMGQSDRIKDKGRNRRTFI-LTFIRIPLVARNECEVAKNV-----VSENNLCAGII 390

QY 334 DSKDSCGSDSGGFHATHYRGTYWLTGIVSMGCGCATGHHGVYTRVSOYIEMLOKLMRS 393

DB 391 GDTRDACDGDSDSGGFVYVFFRGITWFLVGLVSWEGCGATTNNTGITYKVGSYLKWHSYIGE 450

QY 394 EPRPGVLLRA 403

DB 451 K---GVSLKS 457

Search completed: July 1, 2003, 17:53:45
Job time : 83 secs